Research Institute for Quantitative Studies in Economics and Population
Faculty of Social Sciences, McMaster University
Hamilton, Ontario, Canada
L8S 4M4

LINEAR PUBLIC GOODS EXPERIMENTS: A META-ANALYSIS

Jennifer Zelmer

QSEP Research Report No. 361

June 2001

The author is a QSEP Research Associate and Director, Health Reports and Analysis, Canadian Institute for Health Information.

This report is cross-listed as No. 46 in the McMaster University SEDAP Research Paper Series.

Linear Public Goods Experiments: A Meta-Analysis

Jennifer Zelmer

Acknowledgements: Special thanks are due to Stuart Mestelman and Andy Muller (both with the Department of Economics, McMaster University) for their advice and encouragement throughout the development of this paper.

Abstract

The objective of this paper is to use meta-analysis techniques to assess the impact of various factors on the extent of cooperation in standard linear public goods experiments using the voluntary contributions mechanism. Potentially relevant experiments were identified through searches of EconLit, the Internet Documents in Economics Access Service (IDEAS), and a survey article. A total of 349 potentially relevant studies were identified. Of these, 28 (representing a total of 711 groups of participants) met the inclusion criteria. Data were abstracted from these studies using a standardized protocol. Results were analyzed using weighted ordinary least squares. Average group efficiency was the dependent variable.

The major results are that: (1) The marginal per capita return, communication, constant group composition over the session ("partners"), positive framing, and the use of children as subjects had a positive and significant effect ($p<0.05$) on the average level of contribution to the public good; and (2) Heterogeneous endowments to subjects, experienced participants, and soliciting subjects' beliefs regarding other participants' behaviour prior to the start of the session/period had a negative and significant effect. (A number of other factors were not identified as significant.) The meta-analysis results parallel several key findings from previous literature reviews. In addition, they offer parameter estimates and an analysis of significance based on the totality of the available research evidence. More consistent reporting of the results of experiments would greatly improve the ability to conduct this type of research.

Keywords

Public goods, voluntary contribution, experiment

Introduction

Every day, public and private decision-makers make choices with far-reaching consequences. The potential knowledge base that they can draw on is large and expanding rapidly. For instance, academic literature is growing exponentially, doubling every ten to fifteen years (De Solla Price, 1981). In economics alone, EconLit (the electronic bibliographic index maintained by the American Economic Association) now tracks over 600 journals, as well as a wide range of books and dissertations (AEA, 1999). Unfortunately, single experiments or studies in the social sciences rarely provide definitive answers to research questions (Wolf, 1986). Often, many studies scattered throughout the formal, informal, and even unpublished literature can be relevant.

One area where a diverse, sometimes conflicting, body of literature has accumulated is how different institutions, organizational arrangements, and other factors affect individuals' willingness to contribute to public goods. In linear public goods environments, payoff maximizers have a dominant strategy to either contribute all of their tokens or none of their tokens to a group activity. Classical economic theory predicts that markets will not voluntarily achieve optimal allocations of public goods, such as clean air or public television, in this case. For more than two decades, experimental economists and others have endeavoured to test this prediction and factors that influence cooperation in laboratory environments.

Excellent survey articles have been prepared, but as the number of studies grows, the accuracy of this technique to accurately represent the results of the primary research has been questioned. In addition, while the surveys have indicated the direction of effect of a range of variables (e.g. communication and number of subjects), they do not provide estimates of the effect size based on the totality of the evidence. Meta-analysis has the potential to address some of these concerns.

This paper begins with a descriptive overview of meta-analysis as a research methodology, then proceeds to introduce the voluntary contribution mechanism (VCM). Results from a meta-analysis of VCM linear public goods experiments follow.

Meta-Analysis: A Quantitative Research Methodology

The overwhelming volume of current and emerging research literature has led to the development of a range of techniques for synthesizing information (Plath, 1992). One of the earliest and still most commonly used methods for the aggregation and synthesis of knowledge is the literature review. Literature surveys have a long and distinguished history in economics. Some widely read journals, such as the *Journal of Economic Literature and the Journal of Economic Surveys*, have evolved primarily for their dissemination.

Like all secondary analysis, however, the success of literature reviews depends on the quality of both the original research and the subsequent analysis. For instance, as the number of primary studies becomes large, the accuracy of this technique in reproducing the results of the primary research is questionable (Hunter, 1982; Glass, 1981; and Rosenthal, 1991). Another potential problem with classic literature reviews is their subjectivity. There is a risk of bias or error because of selective inclusion of studies, misleading interpretation of study findings, and other factors. For instance, it has been shown that the frequency of citation of clinical trials in the health field is related to their outcome. Studies that reflect prevailing opinion are quoted more frequently than others are (Ravnskov, 1992). In addition, although literature reviews summarize what is known, they provide little forecasting power. A review may, for instance, provide a table with price elasticities found in the different studies surveyed. It will generally not provide an estimated elasticity based on the totality of the available research.

More recently, several other qualitative and quantitative methods have been developed in an attempt to overcome some of the limitations of traditional literature reviews. These newer methodologies include group judgement, consensus conferences, cost effectiveness analysis, mathematical modeling, and meta-analysis (Deber, 1989).

The Development of Meta-Analysis

Methods for summarizing research quantitatively have been discussed for many years. In the 1930s, Sir Ronald Fisher and Karl Pearson were among the first researchers to report the use of formal statistical techniques to combine data from different samples (Wachter and Straf, 1990 and Egger and Smith, 1997). Over time, quantitative synthesis techniques have spread from agriculture and the social sciences, where Glass coined the term meta-analysis in 1976, into other fields, such as medicine.

Meta-analysis is "the statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings into a single, generalizable finding" (Lynn, 1989). Its aims and functions include (Deber, 1989; Plath, 1992; Wolf, 1986; and Jenicek, 1989):

- Summarizing systematically what is known, how it became known, and the quality of this knowledge in a way that draws as much information as possible from existing studies;
- Data from existing studies to be pooled, thereby increasing the ability to identify significant differences between treatment and control groups and to produce better estimates of effect sizes;
- Discovering underlying trends and principles from the accumulation and refinement of a large body of studies;
- Identifying gaps in knowledge, highlighting well investigated areas, and uncovering flaws in existing research;

- Predicting future results by estimating the effect of various values for the independent variables; and

- Developing new hypotheses for future research.

Meta-analysis differs from traditional qualitative literature reviews in a variety of ways (Plath, 1992). First, it systematically identifies as complete a sample of studies pertaining to the issue of interest as possible, rather than highlighting the findings of key research only. The features and results of these studies are then described in a consistent quantitative or quasi-quantitative way. Statistical techniques are subsequently applied to aggregate the findings across studies and objectively examine the relationship between study characteristics and outcomes. Finally, there is a systematic and detailed description of the method used to integrate study results – ensuring replicability of the findings.

The advantages of meta-analysis are well documented (e.g. Plath, 1992; Hedges, 1990; Rosenthal, 1990; Egger and Smith, 1997). Ultimately, meta-analyses should, by systematic cumulation of knowledge, facilitate understanding of a complex body of existing empirical findings. Technical advantages include completeness, explicitness and therefore replicability, a decrease in type II error (several studies with p values of 0.06 are stronger evidence against the null than 1 study with a p value of 0.05), and a better ability to identify and evaluate the impact of moderator variables. Other less direct benefits cited include:

- Influencing the quality of primary studies and the standardized reporting of results;

- Forcing the principal investigators to take an active approach to the literature (not simply reading abstracts and conclusions);

- Enabling better allocation of scarce research funds by objectively identifying gaps in knowledge and well-developed areas of research; and

- Introducing methodological rigour, thereby highlighting the poor quality of previous reviewing practices.

Nonetheless, as when conducting other forms of research synthesis, researchers undertaking meta-analysis must be aware of potential methodological issues. These include potential biases due to subjectivity, the heterogeneity of the component studies, publication bias, variations in the quality and reporting of primary research, the potential interdependence of primary studies, and challenges in synthesizing non-experimental research.

Conducting Meta-Analyses

The structure of a meta-analysis parallels the scientific method: "problem selection, hypothesis formulation, definition and measurement of constructs and variables, sampling, and data analysis" (Glass, 1981). Like any research project, a detailed study protocol is required. Standardized guidelines for reporting results of meta-analyses have also been developed (Moher, et al., 1999)[1].

Once a topic of interest and hypothesis are identified, the data which will be abstracted from primary studies are defined. These include "study" variables, independent and dependent variables linked to the research hypothesis, and "model" variables which describe the primary research studies (Graney, 1990).

For example, in mathematical terms, the functional form for a meta-analysis of voluntary contributions to a public good might be:

$Y = f(S, M) + \mu$, where:

Y = outcome of interest (e.g. level of contributions to the public good)

S = an array of study variables (e.g. the type of subjects and their level of experience)

M = an array of model variables (e.g. the date of publication and an index representing the overall quality of the original studies)

After variables of interest are defined, inclusion/exclusion criteria for studies and a search strategy must be developed (Wolf, 1986). Based on this strategy, an extensive search of the literature is then performed to locate primary research studies. Next, information on study and model variables for each primary sources meeting the inclusion criteria are extracted and coded using a standard template, preferably by at least two separate individuals. Blinding coders to the authors, institutions, journals, sources of funding, and acknowledgements has been reported to improve consistency of results (Egger, Smith, and Phillips, 1997). A range of specialized software exists to assist with this and subsequent stages of the process.

Numerous statistical techniques exist for converting the data from primary studies into a common metric (a single composite measure representing the overall result of the meta-analysis). There is no clear agreement among experts as to what is the best method to use (Graney, 1990; Deber, 1989). A brief summary of possible analytical options is available from the author on request.

While a summary metric is important, meta-analysis should also include a variety of additional analyses. For example, the effects of multiple independent and dependent variables can be examined. Statistical techniques and graphical presentation methods have also been developed to review the distribution of outcomes, identify outliers, and test the heterogeneity of results. Sensitivity analysis to assess the robustness of the combined

estimates to different assumptions and inclusion criteria should also be conducted (Egger, Smith, and Phillips, 1997).

In some cases, the meta-analysis becomes an on-going process. In cumulative meta-analyses, the synthesis is repeated whenever a new study becomes eligible for inclusion (Egger and Smith, 1997). This is becoming an increasingly common practice in the health field.

Applications in Economics

Classical meta-analysis was designed for the assessment of the effectiveness of specific interventions. Researchers have since adapted it to address a wide range of questions, including: 'Does psychotherapy work and, if so, what form works best?', 'Does spending federal money on education really improve student performance', 'Can a single enzyme significantly decrease the risk of heart attack?', and 'Do boot camps reduce juvenile delinquency?' (Hunt, 1997).

While an article on meta-analysis appeared in the *Economist* as early as 1991, applications in economics continue to be relatively rare (van den Bergh, et al., 1997). A recent review identified 19 major studies in environmental, regional, urban, and transport economics (van den Bergh, et al., 1997). Subjects included urban pollution valuation, recreational benefits, recreational fishing, valuation of life estimates, contingent valuation versus revealed preference, noise nuisance, congestion, internal validity of contingent valuation and visibility improvement, multiplier effects of tourism, transport issues, and the price elasticity of demand in travel cost method studies. More recent studies cover areas such as Ricardian equivalence (Stanley, 1998), factors that systematically affect price and income elasticity estimates for gasoline demand (Espey, 1998), more work on regional tourist multipliers (Baaijens, 1998), and threshold public goods (Croson and Marks, 2000).

One of the reasons for the relatively limited take-up of meta-analysis in economics is likely that the emergence of experimental economics is relatively recent (Friedman and Sunder,

1994). Most empirical research in the discipline is based on quasi-experimental/observational studies. The applicability of meta-analysis in this context is controversial. (Spitzer, 1991).

As the body of experimental evidence in economics grows, it is likely that meta-analysis will become increasingly popular in economics. One of the pre-conditions for secondary analysis is a sufficient body of primary research that is relatively consistently reported. In recent years, an increasing number of laboratory experiments have been conducted. Sharing of methodology and investigative tools (e.g. for computer-assisted experiments) has led to more consistent experimental design and reporting (Friedman and Sunder, 1994).

Another factors that may promote interest in meta-analysis is that, due to cost and other logistical factors, economic experiments tend to be relatively small. Some researchers argue that experimental groups, rather than individual results, should be the unit of analysis[2]. With most existing studies, this approach results in small sample sizes and low power. A meta-analysis of similar studies can be more powerful in detecting small differences in results than individual studies alone and the overall results can be generalized more widely (Simes, 1990).

Another advantage of secondary analysis is that it is a relatively inexpensive and well-developed methodology. Access to economic literature, except for unpublished work, is comparatively easy and rapid using electronic databases. In addition, researchers performing meta-analyses do not need access to large numbers of experimental subjects, complex technology, or specialized experimental laboratories. Meta-analysis can also be repeated periodically as new studies emerge. With development of specialized software, systematized searching for relevant primary studies, and standardized reporting of study results, in some disciplines, meta-analysis may "become a field where replication is almost effortless for any interested and capable reader" (O'Rourke, 1989).

As a result, well-conducted meta-analyses based on the results of current experimental economics studies have the potential to effectively and relatively inexpensively summarize a growing body of literature. By objectively synthesizing the existing knowledge base, this type of analysis also has the potential to facilitate transfer of current research evidence to policy-makers and others.

Voluntary Contributions to Public Goods

Roads, parks, police services, projects to improve air quality, and radio — both ancient and modern society abound with examples of public goods, commodities for which use of the good by one agent does not preclude its use by others (Mas-Colell, Whinston, and Green, 1995). With these types of goods, if one person provides a unit of the good, all benefit. An important question, therefore, is under what conditions individuals will voluntarily contribute to the provision of public goods.

For more than two decades, experimental economists have been exploring this issue. A variety of experimental designs have been used. A frequent choice is the voluntary contribution mechanism (VCM) in linear public goods environments.

In these experiments, subjects are divided into groups and play the same game for a finite number of periods. Each period, every subject is endowed with an income of w_i. S/he must then divide this income between a contribution (x_i) to a private account that yields a constant return to themselves only and a contribution to a public account (g_i) where consumption benefits accrue to all group members. At the end of each period, subjects typically learn the aggregate contribution to the public good by all members of their group and their earnings for the period.

Mathematically, individual i's payoff is given by:

$$u_i = \alpha x_i + \beta G$$

S/he must maximize this utility function, subject to a budget constraint ($w_i = x_i+g_i$), a public goods identity ($G = \sum g_i$), and a non-negativity constraint ($g_i \geq 0$). In a linear public goods environment, both α and β are constants.

Typically, experiments are parameterized so that payoff maximizers in finitely repeated games have a dominant strategy to contribute nothing to the public good. The Nash equilibrium strategy is full free-riding. In contrast, the group as a whole is better off when contributions are made to the group account. The Pareto optimal outcome is for all subjects to contribute their entire outcome to the public good.

Over several decades, a rich, but sometimes contradictory, body of experimental evidence has accumulated about this environment. In practice, researchers have found that a significant number of individuals voluntarily contribute to public goods, but that the group optimum is not usually achieved (Ledyard, 1995). A variety of factors appear to influence the extent of cooperation observed. These include characteristics of the experimental environment, such as the marginal per capita return to investments in the public good (MPCR), the number of subjects in a group, repetition, the types of information available to subjects, gender, and homogeneity of endowments and MPCR. "Systemic" factors, including subjects' beliefs, training in economics, experiments with similar types of experiments, feelings of altruism or fairness, and identification with the group, can also be important. So can features of experimental design, such as whether or not communication is allowed between subjects or whether subjects can impose punishments on other group members.

These types of relationships were identified in a recent meta-analysis of threshold (non-linear) public goods experiments (Croson and Marks, 2000). The researchers found that, after controlling for methodological factors like the number of group members, discrete versus continuous contributions, and the amount of the endowment necessary to reach the threshold, there was a positive and significant effect of the step return[3] on the successful provision of the public good.

Study Objective

The objective of this meta-analysis is to synthesize the results of existing experimental evidence on the impact of a variety of organizational arrangements and other factors on the extent of cooperation observed in standard linear public goods experiments using the voluntary contributions mechanism.

Methods

Searching

An extensive search for reports in the economics literature for results of standard single-stage linear public goods experiments using a voluntary contribution mechanism was undertaken. Three sources were used to identify potentially relevant publications and working papers to be screened for possible inclusion in the meta-analysis:

- *EconLit*[4]: An electronic bibliographic index maintained by the American Economics Association that tracks over 600 journals and includes a range of books and dissertations.

- *Internet Documents in Economics Access Service (IDEAS)*[5]: A web-based search engine that includes references to over 60,000 working papers in economics. This site incorporates references from a number of other sources, including BibEc (a web-based bibliography of working papers in economics), WoPec (a web-based service that provides access to downloadable working papers in economics), EconWPA (an economics working paper archive), and ERN (the Economics Research Network).

- References cited in John Ledyard's (1995) classic survey of experimental research related to public goods.

The possibility of accessing an extensive collection of working papers in experimental economics was also explored, but it was not pursued given the current lack of a searchable index. Other options which might be pursued in the future include hand-searching selected journals for relevant experimental results, a review of references cited in papers identified through other search strategies, and a forward search of seminal papers using the Social Sciences Citation Index.

Selection

Candidate studies from these bibliographic sources were included in the meta-analysis if, and only if, they met the following criteria:

- The study consists of a laboratory experiment where observations for one or more sessions were gathered in a controlled environment;
- The study uses a standard voluntary contribution mechanism in a single-stage linear public goods environment with a single public good where the marginal per capita return was less than one[6];
- Group-level results (or averages over groups of similar types) are reported for the experiment for at least one of the outcomes of interest; and
- A report of the experiment's results is obtainable through electronic access or libraries at Universities of Toronto, McMaster University, and/or York University and/or the World Wide Web.

In order to identify studies that might meet these criteria, the following search strategy was used for both EconLit and IDEAS:

- Keyword search for "public goods" and "experiment*"[7]
- Keyword search for "voluntary contribution" and "experiment*"

- Keyword search for “variable contribution” and “experiment*”
- Keyword search for “cooperation” and “experiment*”
- Subject heading search using the Journal of Economic Literature classification system: Subject area H410 (public goods) and keyword “experiment*”
- Subject areas C900, C910, C920, and C990 (design of experiments) and keyword “public goods”

Validity Assessment and Data Abstraction

Data abstracted from primary studies serves as the bridge between what was reported by the primary investigators and what is ultimately reported by a reviewer. In this case, three types of information were recorded: basic bibliographic information for the reference, information about the experiment as a whole, and information about each of the sessions in the experiment. Given the available resources and time, data were abstracted by the author, a graduate student in economics. Two subject experts also reviewed the study selection criteria and abstraction protocols (A. Muller and S. Mestelman) prior to implementation.

All data were entered and stored in a custom-designed Microsoft Access database. This database represents the data source for the quantitative analysis, as well as a historical record of decisions made in the review process. A pilot test of this database and the extraction methodology was conducted using approximately 7 experimental reports.

The database was designed using standard systems analysis techniques. It includes information on the source (e.g. journal article or book) from which the information was abstracted; on the abstraction process; on the experiment as a whole; and on outcome, study, and model variables for each session of the experiment. The data model and data dictionary are available on request from the author. The data dictionary includes standard

coding instructions to improve data reliability and to facilitate potential replication of data extraction and coding.

Quantitative Data Synthesis

Recent meta-analyses in other fields have tended to focus on comparisons and analysis of effect sizes in different studies (Clarke, 1999). To use this measure, information is required on the number of subjects in each group, the mean value for the outcome variable in each group, and the standard deviation of these means. The first two elements were either reported in or could be estimated from material presented in the majority of papers included in this meta-analysis[8]. In contrast, few authors reported standard deviations or the information necessary to calculate them.

As a result, effect size analysis was impractical without requesting additional data directly from the majority of authors. Instead, the meta-regression method used by Croson and Marks (2000) with weighted least squares of group-level results was employed. Two separate analyses were conducted. In the first, the dependent variable was the average efficiency of the group's contributions over the session, weighted by the number of groups represented in the observation[9]. In the second, the dependent variable was the percent decay in efficiency over the session (comparing the rates of contribution to the public good in the first and last periods). Due to significantly less complete reporting in the primary literature, results from this analysis are not reported here but are available from the author on request. An analysis based on the percent of complete free riders (those who contributed none of their endowment to the public good) was also planned. It was not completed because of the small number of studies for which these data were reported by primary authors.

In both cases, the same core independent variables were used. These included study variables that have been hypothesized in the literature to affect cooperation levels, such as the marginal per capita return, group size, gender of subjects[10], extent of subject

experience, and the extent of communication allowed between subjects (Ledyard, 1995). In some cases, where information on these variables was missing from the primary source, standard imputations were used[11]. The complete data abstraction protocol is available from the author on request.

Independent "model" variables describing the study were also used. These ranged from whether subjects were rewarded with cash or other benefits to whether the experiment was run in a fully computerized experimental environment. In addition, while an independent assessment of study quality was not performed, a dummy variable was used to track the effects of situations where quality problems were reported by the primary researchers. Potential publication bias effects were evaluated by the inclusion of a dummy variable indicating whether or not the primary source was a published journal or other article, versus a working paper.[12]

Dummy variables for each experiment were also included where possible.[13] Following Croson and Marks (2000), parameter estimates and other information for these dummy variables are not reported to facilitate interpretation of results. A significant result for one of these dummy variables would suggest that the results for a particular experiment differ systematically from those of other experiments, even after controlling for the study and model variables included in the regression.

Results

Trial Flow

This search strategy yielded 349 primary sources. 243 were from EconLit and 100 were from IDEAS. All six citations from the Ledyard (1995) survey article were also reviewed. Lists of items identified appear in Appendices D, E, and F.

Titles and abstracts for each of these items were screened for conformance with the study inclusion criteria. Where possible, studies not excluded by this preliminary review were then retrieved for a more detailed evaluation.

The results of the search and review process are shown in Figure 1. Common reasons for studies to be dropped include:

- The papers did not include original reports of experiments (e.g. they discussed theoretical considerations or reviewed available literature and previous experimental results);
- The experiments did not use the standard voluntary contribution mechanism or had non-linear payoff functions;
- The experiment discussed was also reported in another publication or working paper that was already included in the meta-analysis (i.e. it was a duplicate report);
- Insufficient information was included in the publication to enable its inclusion in the quantitative data analysis (e.g. key outcome variables such as average contributions were not reported or calculable); and
- The publication was not obtainable through local universities or the World Wide Web (particularly the case for PhD dissertations and working papers not available on the World Wide Web).

INSERT FIGURE 1: STATUS OF POTENTIALLY RELEVANT PRIMARY SOURCES ABOUT HERE

Quantitative Data Synthesis

The meta-analysis focused on the effect of a variety of study and model variables on the average percentage of their endowment that subjects' contributed to the public good. The analysis covered 711 groups (a set of subjects participating in an experimental session

with the same conditions). 13 others were excluded due to missing data for the dependent variable. Results from the weighted least squares regression are shown in Table 1.

INSERT TABLE 1 ABOUT HERE

As expected, several factors significantly affected mean contributions to the public good. For example, as the marginal payoff to a subject from allocations to the public good relative to the private good increased – that is with higher marginal per capita returns – significantly larger contribution levels were observed. Likewise, communication among subjects improved cooperation (even though in game theoretic terms it was "cheap talk"). So did maintaining a constant group of subjects over the session[14], explicitly framing the instructions in terms of altruism or fairness, and using children as subjects. In contrast, experienced subjects tended to make significantly lower contributions than those participating in the experiment for the first time. Two other factors were also identified as leading to lower contributions: providing subjects with heterogeneous endowments and asking them about their beliefs regarding other participants' behavior prior to the start of the session or period.

Variables were also included in the regression in an effort to identify potential influences related to problems with study quality, publication bias, and heterogeneity in the primary literature. Neither of the first two factors proved to be statistically significant, and only one of the individual study dummies (Isaac and Walker, 1988) had $p<0.05$. This suggests that these factors had relatively little overall influence on the meta-analysis results.

Discussion

Meta-analysis is designed to provide a quantitative synthesis of a defined body of literature. Its use continues to be relatively rare in economics, although a handful of studies have begun to appear.

In contrast, qualitative research synthesis is much more established in the field. One of the most cited survey articles on public goods experiments was published by Ledyard in 1995. From a review focusing on six experiments available at the time of writing, he posited a series of stylized facts. Table 3 summarizes Ledyard's findings and those of this meta-analysis for variables that were covered in both studies.

INSERT TABLE 2 ABOUT HERE

Overall, there is substantial correspondence between Ledyard's findings and those from the meta-analysis. To some extent, this provides an external validation of the latter's results. The additional structured abstraction and quantitative analysis in a meta-analysis also allows for formalized hypothesis testing and the generation of parameter estimates based on the totality of the available research evidence. This has the potential to offer significant benefit for policy-makers, as well as for future experimental work.

This type of analysis, however, depends on consistent and complete reporting of the methods and results of primary research. Many of the articles reviewed for this paper – even those published relatively recently – do not include key information, such as the number and gender of subjects and whether communication between subjects was allowed. Research suggests that these are key factors in public goods experiments using the voluntary contribution mechanism.

Application of existing guidelines for reporting the results of economic experiments (e.g. Palfrey and Porter, 1991) has the potential to greatly facilitate both interpretation of primary research and secondary analysis, such as meta-analysis. Sharing experimental data would also be helpful. Both approaches are being strongly encouraged in other fields, such as medicine, where there is increasing interest in meta-analysis (Clarke and Oxman, 1999).

Nevertheless, even with the current state of reporting, additional research might also provide further benefit. One possibility would be to build on the linear public goods meta-

analysis by expanding the search strategy (e.g. to include hand searching of relevant journals) or contacting the original authors to acquire data missing from the published experimental reports. Given the relatively strong predictive power of the current models, however, it is unclear to what extent this additional effort would yield substantive gains.

To expand the scope of the analysis, it would also be possible to conduct additional meta-analyses that covered parallel experimental environments (e.g. common property resources) or incorporated non-linear public goods experiments, building on the work of Croson and Marks (2000). A third possibility would be to leverage the meta-analysis results to conduct further experiments. These could be in areas where current evidence is relatively weak (e.g. cross-cultural comparisons of cooperation outside of America and Western Europe) and/or in an effort to validate specific predictions and parameter estimates from the meta-regression models. The advantage of this approach is that it accepts the strength of the body of existing literature in some areas and focuses further research on areas where outstanding questions remain.

References

American Economics Association. (September 26, 1999). http://www.econlit.org/econlit/ellistjn.html.

Baaijens, SR, Nijkamp, P, and van Montfort, K. (1998). "Explanatory Meta-Analysis for Comparison and Transfer of Regional Tourist Income Multipliers." *Regional Studies.* 32, 9, 839-49.

Clarke, M and Oxman, AD (Eds.). (1999). Cochrane Reviewers' Handbook 4.0 [Updated July 1999]. In Review Manager (RevMan) [Computer Program]. Oxford: The Cochrane Collaboration.

Croson, RTA and Marks, MB. (2000). "Step Returns in Threshold Public Goods: A Meta- and Experimental Analysis." *Experimental Economics*. 2, 239-59.

Deber, R. (1989). "Choices in Methods of Technology Assessment and Their Impact on Decision Making." *Symposium on the Assessment of Health Care Technologies*. Québec: Canadian Intergovernmental Conference Secretariat.

De Solla Price, D. (1981). The Development and Structure of the Biomedical Literature." In Warren, K. (ed.). Coping with the Biomedical Literature. New York: Praeger Publishers. Cited in Siegel, ER, et. al. (1990). "Bibliographic-Retrieval Systems." In Shortliffe, EH, et. al. (eds.). Medical Informatics: Computer Applications in Health Care. Reading, Massachusetts: Addison-Wesley.

Egger, M and Smith, GD. (1997). "Meta-Analysis: Potentials and Promise." *British Medical Journal*. 315, 7119, 1371-4.

Egger, M, Smith, GD, and Phillips, AN. (1997). “Meta-Analysis: Principles and Procedures.” *British Medical Journal.* 315, 7121, 1533-7.

Espey, M. (1998). “Gasoline Demand Revisited: An International Meta-Analysis of Elasticities.” *Energy Economics*. 20, 3, 273-295.

Friedman, D and Sunder, S. (1994). *Experimental Methods: A Primer for Economists*. New York: Cambridge University Press.

Glass, GV, McGaw, B and Smith, ML. (1981). *Meta-Analysis in Social Research*. Beverly Hills: Sage Publications.

Graney, MJ and Engle, VF. (1990). "Meta-Analysis Techniques." *Journal of Gerontological Nursing*. 16, 9, 16-19.

Hedges, LV. (1990). “Directions for Future Methodology.” In Wachter and Straf, 1990.

Hunt, M. (1997). *How Science Takes Stock: The Story of Meta-Analysis*. New York: Russell Sage Foundation.

Hunter, JE, et al. (1982). *Meta Analysis: Cumulating Research Findings Across Studies*. Studying Organizations: Innovations in Methodology Number 4. Beverly Hills: Sage Publications.

Jenicek, M. (1989). "Meta-Analysis in Medicine: Where We Are and Where We Want To Go." *Journal of Clinical Epidemiology*. 42, 1, 35-44.

Ledyard, JO. (1995). “Public Goods: A Survey of Experimental Research.” In Kagel, J and Roth, A. (Eds.). *The Handbook of Experimental Economics*. Princeton: Princeton University Press.

Lynn, MR. (1989). "Meta-Analysis: Appropriate Tool for the Integration of Nursing Research?" *Nursing Research.* 38, 5, 302-305.

Mas-Colell, A, Whinston, MD, and Green, JR. (1995). *Microeconomic Theory*. New York: Oxford University Press.

Moher, D, et al. For the QUORUM Group. (1999). “Improving the Quality of Reports of Meta-Analyses of Randomised Controlled Trials: The QUORUM Statement.” *British Medical Journal*. 354, 9193.

O'Rourke, K and Detsky, AS. (1989). "Meta-Analysis in Medical Research: Strong Encouragement for Higher Quality in Individual Research Efforts." *Journal of Clinical Epidemiology*. 42, 10, 1021-1024.

Palfrey, T and Porter, R. (1991). “Guidelines for Submission of Manuscripts on Experimental Economics." *Econometrica*. 59, 4, 1197-8.

Plath, I. (1992). *Understanding Meta-Analysis: A Consumer’s Guide to Aims, Problems, Evaluation, and Developments*. Baden-Baden: Nomos Verlagsgellschaft.

Ravnskov, U. (1992). “Cholesterol Lowering Trials in Coronary Heart Disease: Frequency of Citation and Outcome.” *British Medical Journal*. 305, 15-9.

Rosenthal, R. (1990). “An Evaluation of Procedures and Results.” In Wachter and Straf, 1990.

Rosenthal, R. (1991). "Meta-Analysis: A Review." *Psychosomatic Medicine*. 53, 3, 247-71.

Simes, J. (1990). "Meta-Analysis: Its Importance in Cost-Effectiveness Studies." *Medical Journal of Australia*. 153, S13-S16.

Spitzer, WO. (1991). "Editorial - Meta-Meta Analysis: Unanswered Questions About Aggregating Data." *Journal of Clinical Epidemiology*. 44, 2, 103-107.

Stanley, TD. (1998). “New Wine in Old Bottles: A Meta-Analysis of Ricardian Equivalence.” *Southern Economic Journal*. 64, 3, 713-27.

Van den Bergh, JCJM, Button, KJ, Nijkamp, P and Pepping, GC. (1997). *Meta-Analysis in Environmental Economics*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Wachter, K W and Straf, ML (Eds.). (1990). *The Future of Meta-Analysis*. New York: Russell Sage Foundation.

Wolf, FM. (1986). *Meta-Analysis: Quantitative Methods for Research Synthesis*. Beverly Hills: SAGE Publications.

References for Studies Included in the Meta-Analysis

Andreoni, J. (1988). "Why Free Ride? Strategies and Learning in Public Goods Experiments." *Journal of Public Economics*. 37, 3, 291-304.

Andreoni, J. (1995). "Cooperation in Public Goods Experiments: Kindness or Confusion?" *American Economic Review*. 85, 4, 891-904.

Andreoni, J. (1995). "Warm-Glow versus Cold-Pickle: The Effects of Positive and Negative Framing on Cooperation in Experiments." *Quarterly Journal of Economics*. 110, 1, 1-21.

Asch, P, Gigliotti, G, and Polito, JA. (1993). "Free Riding with Discrete and Continuous Public Goods: Some Experimental Evidence." *Public Choice*. 77, 2, 293-305.

Cason, TN and Khan, FI. (1999). "A Laboratory Study of Voluntary Public Goods Provision with Imperfect Monitoring and Communication." *Journal of Development Economics*. 58, 2, 533-52.

Chu, HY and Wenchuan, L. (1999). "Free Riding and Consistent Conjectures in the Provision of Public Goods: An Experimental Study." *International Journal of Game Theory*. 28, 1, 89-109.

Croson, RTA. (1996). "Partners and Strangers Revisited." *Economics Letters*. 53, 1, 25-32.

Croson, RTA. (2000). "Thinking Like a Game Theorist: Factors Affecting the Frequency of Equilibrium Play." *Journal of Economic Behavior and Organization*. 41, 3, 299-314.

Dorsey, RE. (1992). "The Voluntary Contributions Mechanism with Real Time Revisions." *Public Choice*. 73, 3, 261-82.

Falkinger, F, Fehr, E, Gachter, S, and Winter-Ebmer, R. (1999). "A Simple Mechanism for the Efficient Provision of Public Goods – Experimental Evidence." Working Paper No. 3. Zurich: Institute for Empirical Research in Economics, University of Zurich.

Fisher, J, et al. (1995). "Heterogeneous Demand for Public Goods: Behavior in the Voluntary Contributions Mechanism." *Public Choice*. 85, 3-4, 249-66.

Gachter, S, Fehr, E, and Kment, C. (1996). "Does Social Exchange Increase Voluntary Cooperation?" *Kyklos*. 49, 4, 541-54.

Goeree, H, Holt, CA, and Laury, SK. (1999). *Altruism and Noisy Behavior in One-Shot Public Goods Experiments*. Paper available through Internet Documents in Economics Access Service.

Harbaugh, WT and Krause, K. (2000). "Children's Altruism in Public Good and Dictator Experiments." *Economic Inquiry*. 38, 1, 95-109.

Isaac, RM and Walker, JM. (1988). "Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism." *Economic Inquiry*. 26, 4, 585-608.

Isaac, RM and Walker, JM. (1988). "Group Size Effects in Public Goods Provision: The Voluntary Contribution Mechanism." *Quarterly Journal of Economics*. 103, 1, 179-99.

Isaac, RM and Walker, JM. (1998). "Nash as an Organizing Principle in the Voluntary Provision of Public Goods: Experimental Evidence." *Experimental Economics*. 87, 5, 829-46.

Isaac, RM, Walker, JM, and Thomas, SH. (1984). "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations." *Public Choice*. 43, 113-149.

Isaac, RM, Walker, JM, and Williams, AW. (1994). “Group Size and the Voluntary Provision of Public Goods: Experimental Evidence Utilizing Large Groups.” *Journal of Public Economics*. 54, 1, 1-36.

Keser, C and van Winden, F. (1996). *Partners Contribute More to Public Goods than Strangers: Conditional Cooperation*. Paper available through Internet Documents in Economics Access Service.

Laury, SK, Walker, JM, and Williams, AW. (1995). “Anonymity and the Voluntary Provision of Public Goods.” *Journal of Economic Behavior and Organization*. 27, 3, 365-80.

McCorkle, S and Watts, M. (1996). “Free Riding Indexes for Ukrainian Economics Teachers.” *Journal of Economic Education*. 27, 3, 233-37.

Nowell, C and Tinkler, S. (1994). “The Influence of Gender on the Provision of a Public Good.” *Journal of Economic Behavior and Organization*. 25, 1, 25-36.

Ockenfels, A and Weimann, J. (1999). ”Types and Patterns: An Experimental East-West German Comparison of Cooperation and Solidarity.” *Journal of Public Economics*. 71, 2, 275-87.

Saijo, T and Nakamura, H. (1995). “The ‘Spite’ Dilemma in Voluntary Contribution Mechanism Experiments.” *Journal of Conflict Resolution*. 39, 3, 535-60.

Weimann, J. (1994). “Individual Behavior in a Free-Riding Experiment.” *Journal of Public Economics*. 54, 1, 1-36.

Wilson, RK and Sell, J. (1997). “’Liar, Liar…’: Cheap Talk and Reputation in Related Public Goods Settings.” *Journal of Conflict Resolution*. 41, 5, 695-717.

Figure Caption

1. Status of Potentially Relevant Primary Sources

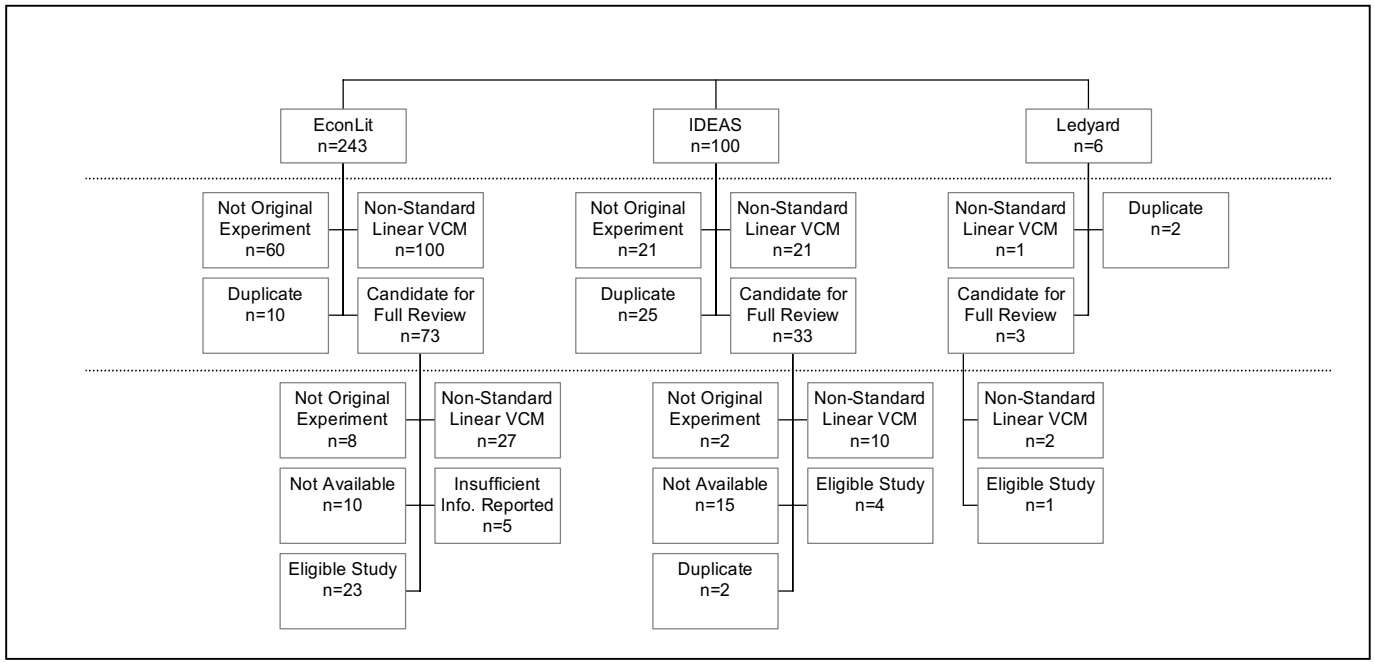

Table 1: Average contributions as a percent of the total endowment[15]

(adjusted r^2=0.6115; observations or sum weights = 711)

Variable	Estimate	Std Error	p-value	sig
Intercept	-14.87	18.44	0.4207	
# Periods	-0.44	0.29	0.1376	
Friendship among Subjects	1.50	7.05	0.8320	
Group Size	0.15	0.09	0.0948	
Cash Rewards	15.36	10.92	0.1605	
Fully Computerized Environment	1.10	5.72	0.8479	
Marginal Per Capita Return	39.53	6.12	<.0001	**
Male Subjects	1.00	13.16	0.9395	
Female Subjects	8.00	12.60	0.5260	
Child Subjects	44.85	22.49	0.0472	*
Heterogeneous MPCR	-0.54	12.65	0.9657	
Heterogeneous Endowments	-14.51	7.10	0.0421	*
Experienced Subjects	-6.15	2.55	0.0167	*
Communication Allowed	40.46	4.16	<.0001	**
Punishment of Subjects Allowed	1.86	6.16	0.7637	
Economics Training	6.05	5.87	0.3039	
Positive Framing	19.30	7.90	0.0151	*
Optimum Announced	-0.46	12.99	0.9716	
End of Session Announced	6.48	9.98	0.5168	
Quality Problems Identified	-5.55	6.95	0.4255	
Imperfect Monitoring of Group Contributions	2.25	7.19	0.7550	
Beliefs re: Others' Behaviour Solicited	-20.00	8.49	0.0193	*
Constant Groups for Session (“partners”)	15.67	3.54	<.0001	**
Subjects from Western Europe	-0.55	10.14	0.9568	
Subjects from Eastern Europe	-10.78	11.20	0.3368	
Japanese Subjects	-10.60	13.22	0.4232	
Published in Journal	-6.28	15.24	0.6807	

Table 2: Stylized Facts from Ledyard (1995) vs. Meta-Analysis Findings

Factor	Ledyard[16]	Meta-Analysis (effect on average contributions)
Marginal per capita return	++	+ ($p<0.01$)
Numbers	00	Not sig.
Repetition	--	Complex[17]
Gender	0	Not sig.
Homogeneity of preferences and endowments	+	Homogeneous MPCR: not sig. Homogeneous Endowments: + ($p<0.05$)
Economics training	--	Not sig.[18]
Experience	--	- ($p<0.05$)
Friendship/group identification	+	Complex[19]
Communication	++	+ ($p<0.01$)
Moral suasion	?	+ ($p<0.05$) for positive framing

Footnotes

1. The meta-analysis results reported in this paper conform to these recommendations.

2. Except in the first period, individual contributions of groups members are not independent because of their common history. But groups' average contributions are independent and therefore are often deemed by many researchers to be the appropriate unit of analysis (Falkinger, 1999).

3. Step return (SR) is a concept developed by the authors that parallels the marginal per capita return in linear public goods environments.

4. This search was conducted on June 28, 2000 using the on-line access to EconLit available at the University of Toronto. No restrictions were placed on the publication type, language of publication, or date of publication.

5. This search was conducted on June 17, 2000. The top 100 items from this search were reviewed (based on a computer-based algorithm to identify the "similarity" of items to the search strategy). The "similarity" score of items retrieved ranged from 96% to 70%. No items with a score of 70-73% met the study inclusion criteria.

6. Various authors recommend that the focus of a meta-analysis be relatively narrow to reduce the heterogeneity of primary studies. For this reason, this limitation has been applied. Future research could also include studies from parallel environments, such as common property resource and public "bads" experiments, non-linear public goods studies, and/or environments where token allocations from one period can be carried forward and invested in subsequent periods.

7. The asterisk (*) serves as a wildcard character. That is, this search would return abstracts with any of the following terms: experiment, experimental, experimentally, etc.

8. The type and level of detail of data reported varied significantly across papers. In many cases, averages over several groups of a similar type were presented graphically and data were abstracted by estimating values from the graphs. In others, regression results at either the individual or group level were reported. Only a few papers included numeric session-by-session data at either the group or individual level.

9. Frequently, group-specific results were not reported. Rather, data was presented for averages over groups of a similar type.

10. Separate dummy variables were used for groups composed of only men or only women. The intercept refers to sessions in which the sex of the subjects was not specified or the group contained both men and women.

11. In spite of the existence of widely available guidelines regarding reporting of experimental results (e.g. Palfrey and Porter, 1991), a large number of primary sources failed to include key information regarding their experimental design. Features such as the marginal per capita return (or the information required to calculate it), group size, and whether or not the experiment was fully computerized were almost universally reported. In contrast, many authors did not fully describe the nature of the subject pool (e.g. whether or not subjects had economics training), when and where the experiment was conducted, and average payments to subjects. Standardized imputations were used for some variables (e.g. subjects were assumed to be inexperienced unless otherwise specified). In two cases (average payment to subjects as a proxy for salience of rewards and the year the experiment was

conducted), variables were dropped from the regression analysis because of inconsistent reporting and lack of a reasonable basis on which to make imputations.

12. Publication bias in meta-analyses is often assessed using funnel plots, but this methodology relies on the availability of effect size information. In this case, these data were not available for a large proportion of included studies. The inclusion of a dummy variable for publication status represents an alternative method of analysis. If this variable is significant, it suggests systematic differences between published and unpublished experimental reports which may suggest publication bias.

13. Dummy variables for particular experiments were excluded if perfect multicollinearity would otherwise have resulted (for example, only one study included Japanese subjects – a variable already included in the regression).

14. In the literature, this is frequently referred to as a "partners", instead of a "strangers", design.

15. Following Croson and Marks (2000), the dummy variables for each study have been suppressed for ease of presentation. Data are available from the author on request.

16. According to Ledyard, "+ means increase, 0 means no effect, - means decrease, and ? means that I do not believe these have been measured yet. A double symbol means the effect is strong and apparently replicable. A single symbol, other than ?, means the effect is apparently there but weak and difficult to replicate."

17. The number of periods in the session was not found to be significant on the average contributions over the session, but there was a significant decline in contributions between the first and last periods. This suggests that there may be a non-linear relationship between repetition and contributions, probably at least partly because of end-game effects.

18. In part, this finding may reflect the relatively low levels of training among most subjects defined as “economics-trained”. Most were students in undergraduate economics courses.

19. Previous friendship among subjects did not have a significant effect on contributions, but maintaining the same group throughout the experiment (the “partners” treatment) did have a positive and significant effect.

Number	Title	Author(s)
No. 338:	A Bayesian Approach for Measuring Economies of Scale with Application to Large Canadian Banks	M.W.L. Chan D. Li D.C. Mountain
No. 339:	Economic Costs of Population Aging	F.T. Denton B.G. Spencer
No. 340:	Population Aging and Its Economic Costs: A Survey of the Issues and Evidence	F.T. Denton B.G. Spencer
No. 341:	How Much Help is Exchanged in Families? Towards an Understanding of Discrepant Research Findings	C.J. Rosenthal L.O. Stone
No. 342:	Did Tax Flattening Affect RRSP Contributions?	M.R. Veall
No. 343:	Families as Care-Providers Versus Care-Managers? Gender and Type of Care in a Sample of Employed Canadians	C.J. Rosenthal A. Martin-Matthews
No. 344:	Alternatives for Raising Living Standards	W. Scarth
No. 345:	Mixed Estimation When the Model And/Or Stochastic Restrictions are Nonlinear	F.T. Denton
No. 346:	A Model of Energy Demand in the U.S. Commercial Sector with Declining Rate Schedules	F.T. Denton D.C. Mountain B.G. Spencer
No. 347:	Projections of the Population and Labour Force to 2046: Canada	F.T. Denton C.H. Feaver B.G. Spencer
No. 348:	Projections of the Population and Labour Force to 2046: The Provinces and Territories	F.T. Denton C.H. Feaver B.G. Spencer
No. 349:	Location of Adult Children as an Attraction for Black and White Elderly Migrants in the United States	K.-L. Liaw W.H. Frey J.-P. Lin
No. 350:	The Effects of Drug Subsidies on Out-of-Pocket Prescription Drug Expenditures by Seniors: Regional Evidence from Canada	T.F. Crossley P. Grootendorst S. Korkmaz M.R. Veall

Number	Title	Author(s)
No. 351:	Describing Disability among High and Low Income Status Older Adults in Canada	P. Raina M. Wong L.W. Chambers M. Denton A. Gafni
No. 352:	Some Demographic Consequences of Revising the Definition of 'Old' to Reflect Future Changes in Life Table Probabilities	F.T. Denton B.G. Spencer
No. 353:	The Correlation Between Husband's and Wife's Education: Canada, 1971-1996	L. Magee J. Burbidge L. Robb
No. 354:	The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1988 Tax Flattening in Canada	M.-A. Sillamaa M.R. Veall
No. 355:	Population Change and the Requirements for Physicians: The Case of Ontario	F.T. Denton A. Gafni B.G. Spencer
No. 356:	2 ½ Proposals to Save Social Security	D. Fretz M.R. Veall
No. 357:	The Consequences of Caregiving: Does Employment Make A Difference?	C.L. Kemp C.J. Rosenthal
No. 358:	Exploring the Effects of Population Change on the Costs of Physician Services	F.T. Denton A. Gafni B.G. Spencer
No. 359:	Reflexive Planning for Later Life: A Conceptual Model and Evidence from Canada	M.A. Denton S. French A. Gafni A. Joshi C. Rosenthal S. Webb
No. 360:	Time Series Properties and Stochastic Forecasts: Some Econometrics of Mortality from The Canadian Laboratory	F.T. Denton C.H. Feaver B.G. Spencer
No. 361:	Linear Public Goods Experiments: A Meta-Analysis	J. Zelmer